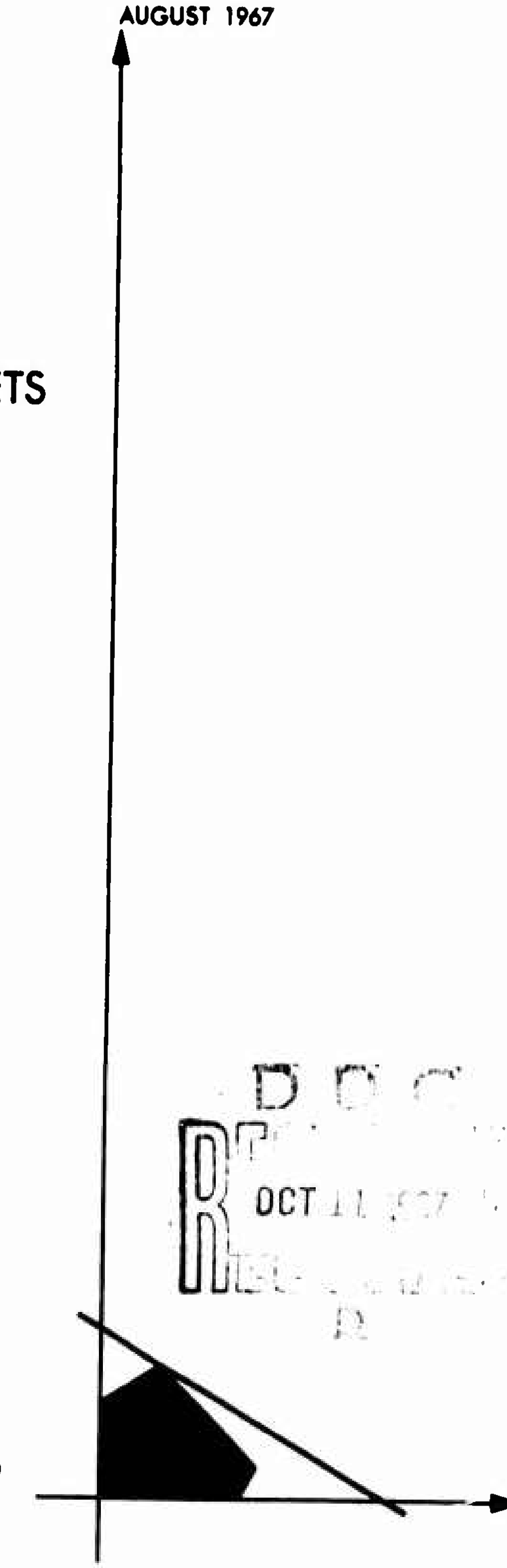
ORC 67-29
AUGUST 1967

PROJECTIONS OF CONVEX POLYHEDRAL SETS

by

David A. Kohler



OPERATIONS RESEARCH CENTER

COLLEGE OF ENGINEERING

UNIVERSITY OF CALIFORNIA - BERKELEY 54

PROJECTIONS OF CONVEX POLYHEDRAL SETS

by

David A. Kohler
Operations Research Center
University of California, Berkeley

August 1967 ORC 67-29

This research has been partially supported by the Office of Naval Research under Contract Nonr-222(83), and the National Science Foundation under Grant GK-1684 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ACKNOWLEDGEMENT

I would like to express my appreciation to the members of my Thesis Committee, Professors R. W. Shephard, David Gale and O. L. Mangasarian for their encouragement. Also, Dr. Roger Wets read the manuscript very carefully and made a number of very helpful comments. I was more grateful for this than even he may realize.

Anyone who reads this thesis will recognize the debt I owe to Noreen Comotto who typed the manuscript.

Finally, but by no means least, I would like to thank my wife [redacted] for her constant encouragement, understanding and patience.

ABSTRACT

The main problem considered is: Given a set of linear inequalities

(1.1) $$Ax + By \geqq d \ ,$$

which defines a set of $(x;y)$, find and concisely define a set Y of y such that if $(x;y)$ solves (1.1) then y belongs to Y and, conversely, if y belongs to Y then there exists an x such that $(x;y)$ solves (1.1).

The solution to this problem involves finding the set of all extreme rays of the convex cone

$$wA = 0 \ , \ w \geqq 0$$

and a method is given for this. The method is compared with other methods for finding extreme rays and points and finally some practical applications are given.

PROJECTIONS OF CONVEX POLYHEDRAL SETS

by

David Kohler

1. INTRODUCTION

Our main concern will be with the problem:

Given a set of linear inequalities

$$Ax + By \geq d \ , \tag{1.1}$$

which defines a set of $(x;y)$, find and concisely define a set Y of y such that if $(x;y)$ solves (1.1) then y belongs to Y and, conversely, if y belongs to Y then there exists an x such that $(x;y)$ solves (1.1). (A, B and d are assumed to be real arrays with dimensions A^{mxn}, B^{mxp}, d^{mx1} and m, n and p are assumed to be finite.)

If we find the set Y then we say that we have "eliminated" x from the system (1.1).

The title of this thesis reflects the fact that a convex polyhedral set is by definition a set describable by a finite system of linear inequalities and the elimination of x amounts in effect to the projection of a set in $(x;y)$-space onto y-space. Practical problems which reduce themselves to the above problem abound in operations research and in other fields, and we will discuss some of them in Section 8.

We shall develop the theory by proving in Section 3 that the set Y can be defined by the system consisting of all of those inequalities in y which are nonnegative linear combinations of the inequalities of the original system (1.1). That is to say, Y may be defined by the system of all

inequalities of the form

$$wAx + wBy \geqq wd \;, \qquad w \in C \tag{1.2}$$

where $C = \{w \mid wA = 0 \;, w \geqq 0\}$. If the only element in C is $w \equiv 0$ then y is clearly unrestricted. Otherwise C is infinite--but we prove in Section 4 that there is a finite subset G of C such that Y may be defined by the system of all inequalities of the type

$$wAx + wBy \geqq wd \;, \qquad w \in G \;. \tag{1.3}$$

In Section 5 we give a method for finding G (and simultaneously the system (1.3)) and in Section 6 we give some computational results. It will become clear, as the theory of Sections 3, 4 and 5 is developed, that the sets C and G are dependent only on A and that (1.3) is the most concise definition of Y which we can give *if we have no prior knowledge of* B *and* d.

But before we can develop this theory we must describe the Fourier-Motzkin Elimination Method. This has historical interest because it is to my knowledge the only method which is currently available and we must explain why it is not often used. However, the main reason for introducing it so early is because it is most useful as a tool for motivating and proving the theory of Sections 3 and 5.

2. THE FOURIER-MOTZKIN ELIMINATION METHOD

Fourier first discovered this method and described it in 1824 in the paper "Solution d'une Question Particuliere du Calcul des Inégalités" [7]. Later T. S. Motzkin drew attention to the method in his Doctoral Thesis "Beitrage zur Theorie der Linearen Ungleichungen" [11] (this is in German but a translation was made by D. R. Fulkerson [11]). A more accessible account may be found in Dantzig [5].

Essentially the method eliminates the unwanted variables $x_1, x_2, \ldots$ *one by one* until linear inequalities in y alone remain. It is only necessary to describe the elimination of one variable--since the elimination of the other variables is just a repetition of this procedure--and we will select x_1 for this purpose. During the elimination of x_1 the variables $y_1, y_2, \ldots, y_p$ do not play a role which is conceptually different from that of the variables $x_2, x_3, \ldots, x_n$. It complicates the expressions to include them explicitly and so we will ignore the y variables and describe the elimination of x_1 from the system

$$(2.1) \qquad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \geq \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_m \end{bmatrix}$$

We partition the set of row indices as follows

$$I^+ = \{i \mid a_{i1} > 0\}$$
$$I^- = \{i \mid a_{i1} < 0\}$$
$$I^o = \{i \mid a_{i1} = 0\}$$

and examine separately the two cases

(A) I^+ *or* I^- *or* both are empty sets.

(B) Both I^+ and I^- are nonempty.

Case A

Here the elimination of x_1 can be effected very easily, thus:

The set of feasible $(x_2, x_3, \ldots, x_n)$ may be defined by the system of inequalities

$$(2.2) \qquad a_{i2}x_2 + a_{i3}x_3 + \ldots\ldots + a_{in}x_n \geq d_i \,, \qquad \forall \; i \in I^o$$

because if $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ solves (2.1) then it solves (2.2) and, conversely, if $(\bar{x}_2, \bar{x}_3, \ldots, \bar{x}_m)$ solves (2.2) and if $I^+(I^-)$ is nonempty then we can always choose an $\bar{x}_1$ so positively (negatively) large that (2.1) is satisfied by $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \ldots, \bar{x}_n)$.

We draw special attention to the fact that the set of feasible $(x_2, x_3, \ldots, x_n)$ is unrestricted if I^o is empty.

Case B

In this case we proceed first by reformulating the statement of the inequalities (2.1). If

$$a_{i1}x_1 + a_{i2}x_2 + \ldots\ldots + a_{in}x_n \geq d_i$$

is an inequality of (2.1) and if $i \in I^+$ then we reformulate it as

$$x_1 \geq \frac{-a_{i2}x_2}{a_{i1}} \frac{-a_{i3}x_3}{a_{i1}} \cdots \frac{-a_{in}x_n}{a_{i1}} + \frac{d_i}{a_{i1}} \,,$$

and let l_i represent the expression on the R.H.S.; alternatively if $i \in I^-$ we reformulate it as

$$\frac{-a_{i2}x_2}{a_{i1}} \frac{-a_{i3}x_3}{a_{i1}} \cdots \frac{-a_{in}x_n}{a_{i1}} + \frac{d_i}{a_{i1}} \geq x_1$$

and let u_i represent the expression on the L.H.S.

Thus (2.1) may be restated as

$$(2.3) \quad \begin{cases} x_1 \geq l_i , & \forall \ i \in I^+ \\ u_i \geq x_1 , & \forall \ i \in I^- \\ a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{in}x_n \geq d_i , & \forall \ i \in I^o . \end{cases}$$

It is now clear that we can describe the set of feasible $(x_2, x_3, \ldots, x_n)$, after the elimination of x_1, by the following inequalities

$$(2.4) \quad \begin{cases} u_k \geq l_i & , \forall \ i \in I^+ \text{ and } k \in I^- \\ a_{i2}x_2 + a_{i3}x_3 + \ldots + a_{in}x_n \geq d_i & , \forall \ i \in I^o \end{cases}$$

That is, any set of $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ which satisfies (2.1) also satisfies (2.4). Conversely, if $(\bar{x}_2, \ldots, \bar{x}_n)$ solves (2.4) then we can always choose an $\bar{x}_1$ so that $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ solves (2.1). The reason for the latter is that the inequalities (2.4) ensure that

$$\max_{i \in I^+} (l_i) \leq \min_{k \in I^-} (u_k)$$

and so any x_1 in the interval $\left[\max_{i \in I^+} (l_i) , \min_{k \in I^-} (u_k)\right]$ will satisfy

$$x_1 \leq u_k , \qquad \forall \ k \in I^-$$

and

$$l_i \leq x_1 , \qquad \forall \ i \in I^+ .$$

This completes the description of the elimination of x_1 from (2.1). The elimination of x from the system (1.1) simply involves repeating this process for $x_2, x_3, \ldots, x_n$ in turn. (Of course no significance is attached to the *sequence* of elimination of the unwanted variables. Any other sequence of elimination would serve the purpose equally well--but note that different sequences *may* yield different sets of inequalities defining Y.) One last point is worth emphasizing. It is this: if *no* inequalities are generated by the elimination of one of the unwanted variables then this immediately implies that the set Y is unrestricted and that we need not, indeed cannot, perform any subsequent eliminations.

An example will illustrate the general method. We wish to eliminate x_1 and x_2 from the system

$$
\begin{array}{rrrl}
-x_1 + & x_2 + & y_1 & \geq 1 \\
- & x_2 + & 2y_1 & \geq -1 \\
-2x_1 + & 4x_2 + & 3y_1 & \geq 1 \\
- & 3x_2 - & 4y_1 & \geq -1 \\
- & 2x_2 - & 5y_1 & \geq 1 \\
 & 5x_2 - & 6y_1 & \geq -1 \\
 & 4x_2 - & 7y_1 & \geq -2 \\
 & & 8y_1 & \geq 4 \\
 & - & 9y_1 & \geq -2
\end{array}
\tag{2.5}
$$

The first column is all nonpositive. So the elimination of x_1 gives:

$$
\begin{aligned}
(2.6) \qquad & - x_2 + 2y_1 \geqq -1 \\
& - 3x_2 - 4y_1 \geqq -1 \\
& - 2x_2 - 5y_1 \geqq 1 \\
& 5x_2 - 6y_1 \geqq -1 \\
& 4x_2 - 7y_1 \geqq -2 \\
& 8y_1 \geqq 4 \\
& - 9y_1 \geqq -2
\end{aligned}
$$

We reformulate this as

$$
\begin{aligned}
(2.7) \qquad & x_2 \geqq \frac{6}{5} y_1 - \frac{1}{5} \\
& x_2 \geqq \frac{7}{4} y_1 - \frac{2}{4} \\
& 2 y_1 + 1 \geqq x_2 \\
& - \frac{4}{3} y_1 + \frac{1}{3} \geqq x_2 \\
& - \frac{5}{2} y_1 - \frac{1}{2} \geqq x_2 \\
& 8y_1 \geqq 4 \\
& - 9y_1 \geqq -2
\end{aligned}
$$

and the elimination of x_2 gives

$$
\begin{aligned}
(2.8) \qquad & 2 y_1 + 1 \geqq \frac{6}{5} y_1 - \frac{1}{5} \\
& - \frac{4}{3} y_1 + \frac{1}{3} \geqq \frac{6}{5} y_1 - \frac{1}{5} \\
& - \frac{5}{2} y_1 - \frac{1}{2} \geqq \frac{6}{5} y_1 - \frac{1}{5} \\
& 2 y_1 + 1 \geqq \frac{7}{4} y_1 - \frac{2}{4} \\
& - \frac{4}{3} y_1 + \frac{1}{3} \geqq \frac{7}{4} y_1 - \frac{2}{4} \\
& - \frac{5}{2} y_1 - \frac{1}{2} \geqq \frac{7}{4} y_1 - \frac{2}{4} \\
& 8y_1 \geqq 4 \\
& - 9y_1 \geqq -2
\end{aligned}
$$

Or, putting the variables on the L.H.S. and the constant terms on the R.H.S. we have

$$
(2.9)\qquad
\begin{aligned}
\frac{4}{5}\, y_1 &\geq -\frac{6}{5}\\
-\frac{38}{15}\, y_1 &\geq -\frac{8}{15}\\
-\frac{37}{10}\, y_1 &\geq \frac{3}{10}\\
\frac{1}{4}\, y_1 &\geq -\frac{3}{2}\\
-\frac{37}{12}\, y_1 &\geq -\frac{5}{6}\\
-\frac{17}{4}\, y_1 &\geq 0\\
8\, y_1 &\geq 4\\
-9\, y_1 &\geq -2
\end{aligned}
$$

The main problem which occurs when using this method is that Case B seems to occur more frequently than Case A in practical problems. In Case B if m^+ and m^- are the numbers of indices in I^+ and I^- respectively then the number of inequalities after elimination is

$$m^+ \cdot m^- + (m - m^+ - m^-)$$

or an increase of

$$(m^+ - 1)(m^- - 1) - 1 \ .$$

Thus the number of inequalities generated by the elimination of a variable can be, and usually is, much larger than the number before elimination. After a few eliminations the number usually becomes impossibly large, even when m is quite small. For example, if A is the matrix

$$\begin{bmatrix} 1 & 1 & 6 \\ 1 & -4 & 4 \\ 1 & -11 & 10 \\ -1 & 7 & -6 \\ -1 & 3 & -7 \\ -1 & -8 & 5 \\ -1 & -9 & 1 \end{bmatrix}$$

then the elimination of x_1 generates 12 inequalities,

the elimination of x_2 generates 27 inequalities and

the elimination of x_3 generates 126 inequalities.

However, the Fourier-Motzkin Method has other uses. In particular we can make the following deductions from it:

(1) The Fourier-Motzkin Method demonstrates something which many people would regard as intuitively reasonable--namely that the set Y of feasible y can always be defined by a finite set of linear inequalities in y alone; or, in other words, the projection of a convex polyhedral set onto a vector subspace is itself also a convex polyhedral set.

(2) The inequalities, if any, which are generated by an elimination are *nonnegative linear combinations* of the original inequalities. In Case A, of course, this is obvious since the inequalities after elimination *are* some of the original ones. In Case B it can be demonstrated easily by comparing (2.3) with (2.4). Reformulating (2.3) as

$$\text{(2.10)} \quad \begin{cases} x_1 \geq l_i , & \forall\ i \in I^+ \\ -x_1 \geq -u_i , & \forall\ i \in I^- \\ a_{i2}x_2 + a_{i3}x_3 + \ldots\ldots + a_{in}x_n \geq d_i , & \forall\ i \in I^0 \end{cases}$$

we can see that each inequality in (2.4) of the form

$$u_k \geqq l_i$$

is the *sum* of two inequalities in (2.10) and, since the inequalities of (2.10) are *positive* scalar multiples of some of those of (2.1), we have proved our assertion.

This means that the elimination of x_j, if it does generate inequalities, is equivalent to premultiplying by a matrix M^j the system which we had obtained after eliminating x_{j-1}. The system we obtain after eliminating $x_1, x_2, \ldots, x_j$ is therefore

$$M^j \cdot M^{j-1} \cdot \ldots \cdot M^2 \cdot M^1 \cdot Ax + M^j \cdot \ldots \cdot M^2 \cdot M^1 \cdot By$$
$$\geqq M^j \cdot \ldots \cdot M^2 \cdot M^1 \cdot d$$

The matrices M^j have a special form and may be constructed easily and in an obvious manner from the column of coefficients of x_j, the variable about to be eliminated. Let

$$\begin{bmatrix} \bar{a}_{1j} \\ \bar{a}_{2j} \\ \bar{a}_{3j} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix}$$

be the column of coefficients of x_j before elimination. A row r of M^j is constructed from each pair of elements $\bar{a}_{i'j}$ and $\bar{a}_{i''j}$, one of which is negative and the other positive. The row $(m_{r1}, m_{r2}, m_{r3}, \ldots)$ has the form

$$m_{rk} = \begin{cases} |\bar{a}_{i'j}|^{-1} & \text{if} \quad k = i' \\ |\bar{a}_{i''j}|^{-1} & \text{if} \quad k = i'' \\ 0 & \text{otherwise} \end{cases}$$

there is also a row s of M^j corresponding to each element $\bar{a}_{ij}$ which is zero. In this case

$$m_{s\ell} = \begin{cases} 1 & \text{if} \quad \ell = i \\ 0 & \text{otherwise} \end{cases}$$

For example if the coefficients of x_j before elimination are

$$\begin{bmatrix} 2 \\ 3 \\ -4 \\ -5 \\ 0 \\ 0 \end{bmatrix}$$

then

$$M^j = \begin{bmatrix} \frac{1}{2} & & \frac{1}{4} & & & \\ \frac{1}{2} & & & \frac{1}{5} & & \\ & \frac{1}{3} & \frac{1}{4} & & & \\ & \frac{1}{3} & & \frac{1}{5} & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix}$$

We are now in a position to state the following definition and prove the subsequent lemma:

Definition of F

Apply the Fourier-Motzkin Method to the elimination of x from (1.1) and if the method *does* generate inequalities then define F to be the set of all rows of the matrix

$$M^n \cdot M^{n-1} \cdot \ldots\ldots \cdot M^2 \cdot M^1$$

Alternatively if no inequalities are generated define F to be the vector

$$w \equiv 0 .$$

(N.B. F is dependent not only on A but also on the order of the variables eliminated. It is assumed in defining F that this sequence is known.)

Lemma 2.1

Y may be defined by the system

(2.11) $$wAx + wBy \geqq wd , \qquad \forall \; w \in F$$

Proof:

The Fourier-Motzkin Method generates the system (2.11) if it generates any inequalities at all. On the other hand Y is unrestricted and may be defined by the inequality

$$0 \cdot Ax + 0 \cdot By \geqq 0 \cdot d$$

Q.E.D.

3. THE KEY THEOREM

Consider the system of inequalities (1.1).

Let S be the set of all nonnegative linear combinations of the inequalities of (1.1) which annihilate A, i.e., the set of inequalities

$$wAx + wBy \geqq wd , \qquad \forall\ w \in C$$

where $C = \{w \mid wA = 0 ,\ w \geqq 0\}$.

Also, let T be the set of inequalities

$$wAx + wBy \geqq wd , \qquad \forall\ w \in F .$$

As we showed in the previous section F is a subset of C and therefore T is a subset of S.

Now let U be the set of y which satisfies (1.1), i.e., the set of y for which there exists an x such that $(x;y)$ satisfies (1.1); let V be the set of y which satisfies S and finally let W be the set of y which satisfies T.

Then the following statements are true whatever B and d may be:

Any solution of a system of linear inequalities also solves any nonnegative linear combination of these inequalities

$$\therefore\ U \subseteq V .$$

Also $S \supseteq T$

$$\therefore\ V \subseteq W .$$

And the Fourier-Motzkin Method implies that $W \equiv U$. This proves the

Theorem 3.1

Whatever B and d may be, $U \equiv V \equiv W$.

The importance of this theorem should be clear after reading the next sections but we may hint at it now. Although the set S may be infinite, we only require a relatively small number of these inequalities in order to adequately define U . These inequalities are necessarily a subset of T , the inequalities generated by the Fourier-Motzkin Method.

4. THE CONVEX CONE: $wA = 0$, $w \geqq 0$

Let C denote the set of all solutions of

$$wA = 0 \ , \ w \geqq 0 \ . \tag{4.1}$$

It is important to see that if a particular element w^o of C is a nonnegative linear combination of other elements $w^1, w^2, w^3, \ldots$ of C then the corresponding inequality

$$w^o By \geqq w^o d$$

is a redundant inequality in the set S . This is because if

$$w^o = \lambda^1 w^1 + \lambda^2 w^2 + \lambda^3 w^3 + \ldots .$$

for some nonnegative scalars $\lambda^1, \lambda^2, \lambda^3, \ldots$ then

$$w^o B = \lambda^1 (w^1 B) + \lambda^2 (w^2 B) + \ldots$$

and

$$w^o d = \lambda^1 (w^1 d) + \lambda^2 (w^2 d) + \ldots$$

Thus it is natural to search for the smallest subset G of C which possesses the property that every element of C is a nonnegative linear combination of the elements of G . Theorem 4.2 below provides the information we need in order to describe G but before we can state and prove it we need the following notation, definitions and lemma:

(i) Let ξ_k represent the k^{th} row of A . (A has a finite number of rows.)

(ii) If w^o is a given m-vector then we define the subset $I(w^o)$ of the indices $\{1,2,3, \ldots, m\}$ as

$$I(w^o) = \{i \mid w_i^o \neq 0\} .$$

(iii) If w^o is a vector in C we define the subset $K(w^o)$ of rows of A as follows

$$K(w^o) = \{\xi_i \mid i \in I(w^o)\} .$$

(iv) We will use the notation (w^o) to represent the set $\{\lambda w^o \mid \lambda \geq 0\}$.

(v) A set of s rows of A is said to possess r dependent rows if it

(a) contains a set of $(s-r)$ independent rows, and

(b) does not contain a set of $(s-r+1)$ independent rows.

(Note that the row of all zeros alone constitutes a dependent set.)

(vi) If w^o is a nonzero element in C and if $K(w^o)$ contains exactly one dependent row then w^o will be called an *extreme vector* of C and the set (w^o) will be called an *extreme half-line* of C .

(vii) A *sufficient set of extreme vectors* of C is a subset of C consisting of exactly one extreme vector from every one of the extreme half-lines of C .

(viii) Lemma 4.1

Let w^1 be a nonzero solution of $wA = 0$ and let $\{\xi_i \mid i \in I(w^1)\}$ contain exactly one dependent row. If w^2 is another solution of $wA = 0$ and if

$$\{\xi_i \mid i \in I(w^2)\} \subseteq \{\xi_i \mid i \in I(w^1)\}$$

then w^2 is a scalar multiple of w^1 .

Proof:

Consider the system

$$\sum_{i \in I(w^1)} w_i \xi_i = 0 \tag{4.2}$$

where $\{\xi_i \mid i \in I(w^1)\}$ contains one dependent row. If i^o is some i in $I(w^1)$ and if we are given the value of w_{i^o} in a solution of (4.2) then the values of the other variables required to solve (4.2) are precisely determined.

Therefore if $w_{i^o} = w^2_{i^o} = kw^1_{i^o}$ in a solution of (4.2), the other variables must be

$$w_i = w^2_i = kw^1_i \qquad \forall\ i \in I(w^1) .$$

Q.E.D.

Theorem 4.2

If C contains a nonzero element and if G is a sufficient set of extreme vectors of C then G is finite and every element of C is a nonnegative linear combination of the elements of G. Furthermore if G' is any other subset of C which has the property that every element of C is a nonnegative linear combination of the elements of G', then each element of G is a positive scalar multiple of some element of G' and therefore no smaller set than G has this property.

Proof:

Let w^o be any nonzero element of C. $K(w^o)$ must contain at least one redundant row since w^o would necessarily be zero if $K(w^o)$ were an independent set. If $K(w^o)$ contains exactly one dependent row then w^o is

itself a scalar multiple of an element of G by Lemma 4.1. Alternatively, if $K(w^o)$ contains more than one dependent row, we select any row in $K(w^o)$, say ξ_1, and define the subset $I(w^o,1)$ of the indices $\{1,2,3, \ldots, m\}$ as

$$I(w^o,1) = \{i \mid w_i^o > 0 , i \neq 1\}$$

and then form the following system of linear equations:

$$\sum_{i \in I(w^o,1)} v_i \xi_i = -w_1^o \xi_1 . \tag{4.3}$$

(Notice that $\{v_i \mid i \in I(w^o,1)\}$ is a set of variables but w_1^o is a constant.)

Demonstrably (4.3) has a nonnegative solution and the Simplex Method of Linear Programming can be used to show that if there is a nonnegative solution of a system of linear equations then there is at least one nonnegative *basic* solution. Let $\{\bar{v}_i \mid i \in I(w^o,1)\}$ be such a nonnegative basic solution of (4.3) and define $\bar{w}$ as

$$\bar{w}_i = \begin{cases} \bar{v}_i & \text{if} \quad i \in I(w^o,1) \\ w_1^o & \text{if} \quad i = 1 \\ 0 & \text{otherwise} \end{cases}$$

We next find the largest (positive) scalar k such that $(w^o - k\bar{w})$ is nonnegative and we can then form w^o as the sum

$$w^o = (w^o - k\bar{w}) + k\bar{w}$$

where

(i) $\bar{w}$ is a scalar multiple of an element of G, and

(ii) $K(w^o - k\bar{w})$ is a strict subset of $K(w^o)$.

We repeat the above process, using $(w^o - k\bar{w})$ where above we used w^o, until we have expressed w^o as a nonnegative linear combination of elements of G. The phrase "if C contains a nonzero element" is inserted in the statement of the theorem in order to ensure that G be nonempty and when this is the case it is trivially obvious that the zero vector is a nonnegative linear combination of the elements of G. So we have proved that *any* element of C is a nonnegative combination of the elements of G.

G must be finite because there are only finitely many combinations of rows of A which have one dependent row.

Only the last assertion remains to be proved and this depends on the fact that if an element, say w^1, of G is a *positive* linear combination of some other elements $\bar{w}^1, \bar{w}^2, \bar{w}^3, \ldots$ of C then these elements are nonnegative scalar multiples of w^1. We show this as follows:

Let

$$w^1 = \lambda^1\bar{w}^1 + \lambda^2\bar{w}^2 + \lambda^3\bar{w}^3 + \ldots .$$

where $\bar{w}^1, \bar{w}^2, \bar{w}^3, \ldots, \in C$ and $\lambda^1, \lambda^2, \lambda^3, \ldots, > 0$. Since $w^1, \bar{w}^1, \bar{w}^2, \bar{w}^3, \ldots$ are all nonnegative, $K(w^1), K(w^2), K(w^3), \ldots$ must be subsets of $K(w^1)$ because $\bar{w}^i_i > 0 \Rightarrow w^1_i > 0$. But $K(w^1)$ contains exactly one dependent row and so by Lemma 4.1 $\bar{w}^1, \bar{w}^2, \bar{w}^3, \ldots$ must be scalar multiples of w^1 and any set G' which possesses the property described in the statement of the theorem, must therefore contain at least one element from each extreme half-line. Since G possesses only one element from each of these extreme half-lines there can be no subset of C smaller than G with this property.

Q.E.D.

5. A CONCISE DEFINITION OF Y

We now have a solution, in theory at least, to the problem of finding a concise definition of the set Y which we stated in Section 1. We may express the solution as follows:

Given the system

$$Ax + By \geqq d \tag{5.1}$$

we examine the associated system

$$wA = 0 \;, \; w \geqq 0 \;. \tag{5.2}$$

If the only solution of (5.2) is $w \equiv 0$ then the set Y is unrestricted, i.e., for any y whatsoever we can always find an x such that (5.1) is satisfied. Alternatively if (5.2) has more than one solution then the system

$$wAx + wBy \equiv wBy \geqq wd \;, \qquad \forall\, w \in G \tag{5.3}$$

where G is a sufficient set of extreme vectors of C, is a concise definition of Y. In fact if we have no prior knowledge of B and d then there can be no more concise definition of Y than (5.3)--as we can show by letting $B = I$ and $d = 0$. Then (5.3) reduces to

$$wy \geqq 0 \;, \qquad \forall\, w \in G$$

which clearly does not contain a redundant inequality because, by Farkas' Lemma [6], a member of a system of homogeneous linear inequalities is redundant if and only if it is a nonnegative linear combination of the other inequalities of the system and we have deliberately chosen G so that this is not the case.

In this section we will describe an adaptation of the Fourier-Motzkin

Method which determines whether C has more than one solution and, if so, produces a definition of Y in the form (5.3). It makes use of the following theorems:

Theorem 5.1

Let C contain more than one solution and let H be a subset of C such that if $(x;y)$ solves (1.1) then it solves the system

(5.4) $$wAx + wBy \equiv wBy \geqq wd \,, \qquad \forall \;\; w \in H$$

and if y solves (5.4) then there exists an x such that $(x;y)$ solves (1.1), whatever B and d may be.

Then H contains a sufficient set of extreme vectors of the set C .

Proof:

The solution sets of the system (5.4) and of the system

(5.5) $$wAx + wBy \equiv wBy \geqq wd \,, \qquad \forall \;\; w \in G$$

are identical, *whatever* B *and* d *may be.* So we may legitimately allow B and d to take on the special values I and 0 respectively in order to examine the relationship between G and H . Then (5.4) and (5.5) reduce to

(5.6) $$wy \geqq 0 \,, \qquad \forall \;\; w \in H$$

(5.7) $$wy \geqq 0 \,, \qquad \forall \;\; w \in G$$

respectively. Each inequality of (5.7) is satisfied by every solution of the system (5.6) and therefore, by Farkas' Lemma, each element of G is a non-negative linear combination of the elements of H .

The required result then follows as a direct consequence of Theorem 4.2.

Q.E.D.

Theorem 5.1 tells us, for example, that we may restrict our search for a sufficient set of extreme vectors to the set F which we defined at the end of Section 2. The problem is to distinguish them from the other vectors in F . Theorems 5.2 and 5.3 below describe the tools with which we may conveniently do this.

<u>Theorem 5.2</u>

w^o is an extreme vector of C if and only if

(i) w^o is a nonzero element of C , and

(ii) there does *not* exist a nonzero element w^1 of C such that $I(w^1)$ is a strict subset of $I(w^o)$.

<u>Proof</u>:

Assume that w^o is an extreme vector and that w^1 is an element of C such that $I(w^1)$ is a strict subset of $I(w^o)$. Then $\{\xi_i \mid i \in I(w^o)\}$ contains one dependent row and by Lemma 4.1 w^1 is a scalar multiple of w^o , say $w^1 = kw^o$. But $I(w^1)$ is a *strict* subset of $I(w^o)$ and therefore there exists an i , say i' , belonging to $I(w^o)$ but not to $I(w^1)$. This implies that $w^1_{i'} = kw^o_{i'} = 0$ which in turn implies that w^1 must be zero. Therefore if w^o is an extreme vector there cannot be a *nonzero* element w^1 of C such that $I(w^1)$ is a strict subset of $I(w^o)$.

Arguing in the opposite direction, assume that w^o is a nonzero element of C and that there does not exist a nonzero $w^1 \in C$ such that $I(w^1)$ is a strict subset of $I(w^o)$. By Theorem 4.2 every nonzero element of C is a *positive* linear combination of one or more members of some given sufficient set of extreme vectors $\{w^1, w^2, w^3, \ldots\}$

$$w^o = \lambda^1 \bar{w}^1 + \lambda^2 \bar{w}^2 + \lambda^3 \bar{w}^3 + \ldots. , \qquad \lambda^1, \lambda^2, \lambda^3, \ldots > 0$$

where $\bar{w}^1, \bar{w}^2, \bar{w}^3, \ldots$ are elements of $\{w^1, w^2, \ldots\}$. But our assumption, together with the fact that $w^o, \bar{w}^1, \bar{w}^2, \ldots$ are nonnegative, implies that $I(w^o) \equiv I(\bar{w}^1) \equiv I(\bar{w}^2) \equiv \ldots$ which in turn implies, by Lemma 4.1, that $w^o, \bar{w}^1, \bar{w}^2, \ldots$ are scalar multiples of one another. Therefore w^o is an extreme vector (and is in fact a positive scalar multiple of exactly one element of a sufficient set of extreme vectors).

Q.E.D.

This theorem is useful in the following way: We have established that F contains a sufficient set of extreme vectors. So we simply search through F discarding any vector w^1 if there exists another vector w^o in F such that $I(w^o)$ is a strict subset of $I(w^1)$. I have been unable to prove that we can discount the possibility of the Fourier-Motzkin Method generating more than one extreme vector from the same extreme half-line. Should this occur we need only keep one of them.

Theorem 5.3

If A contains n columns then no extreme vector can contain more than $n+1$ nonzero components.

Proof:

If w^o is an extreme vector then $\{\xi_i \mid i \in I(w^o)\}$ contains one dependent row and a set of no more than n independent rows.

Q.E.D.

This means that we may discard those vectors in F which have more than $n+1$ nonzero components. It is a simpler and quicker method than the one derived from Theorem 5.2 but, unlike the latter, it may not discard all of the unwanted vectors in F .

It is important to realize that we do not have to wait until we have eliminated all n variables $x_1, x_2, \ldots, x_n$ by the Fourier-Motzkin Method

before we can apply these methods. They may be applied at any time and as often as we like during the elimination of the variables because, although the theorems were stated with regard to the complete matrix, they are of course equally true when applied to a matrix $A^{(k)}$ containing any k columns of A, for $k = 1,2,3, \ldots, n$ (or any other matrix for that matter).

Now we can describe the method for finding a sufficient set of extreme vectors. It goes as follows:

Beginning with the system

$$Ax + By \geqq d \tag{5.8}$$

we eliminate x_1 using the Fourier-Motzkin Method and if this generates any inequalities at all it is equivalent to premultiplying (5.8) by a matrix M^1, as we showed in Section 2. (If the Fourier-Motzkin Method generates no inequalities at this or a later stage we deduce at once that Y is unrestricted.) We examine the rows of M^1 to see if there are any which do not belong to a sufficient set of extreme vectors of the convex cone

$$(w_1, w_2, \ldots, w_m) \begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ a_{m1} \end{bmatrix} = 0$$

$$w_i \geqq 0 , \qquad i = 1,2, \ldots, m$$

using the criteria of Theorems 5.2 and 5.3. Actually each row of M^1 is a member of a sufficient set of extreme vectors, whatever A may be, because a row of M^1 contains *either* one nonzero element which is the only nonzero element in its column *or* a unique combination of two nonzero elements, and

therefore each row of M^1 is an extreme vector by Theorem 5.2. We define $\bar{M}^1$ to be M^1 and proceed to the elimination in turn of the other variables. The procedure for $x_2, x_3, \ldots$ etc. is the same for each and we will describe it for the general case, x_j :

$$\begin{aligned} &(\bar{M}^{j-1} \cdot \bar{M}^{j-2} \cdot \ldots \cdot \bar{M}^2 \cdot \bar{M}^1) \cdot Ax + \\ (5.9) \quad &(\bar{M}^{j-1} \cdot \ldots \cdot \bar{M}^2 \cdot \bar{M}^1) \cdot By \geqq (\bar{M}^{j-1} \cdot \ldots \cdot \bar{M}^1) \cdot d \end{aligned}$$

is the system we have obtained by eliminating $x_1, x_2, \ldots, x_{j-1}$. Actually it is more convenient not to carry this system explicitly. We only need the matrix $(\bar{M}^{j-1} \cdot \ldots \cdot \bar{M}^2 \cdot \bar{M}^1)$ and the original A , B and d and we can obtain updated columns as we require them. The j^{th} column of (5.9),

$$(\bar{M}^{j-1} \cdot \ldots \cdot \bar{M}^2 \cdot \bar{M}^1) \begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ \cdot \\ a_{mj} \end{bmatrix}$$

is computed and the Fourier-Motzkin Method is applied to eliminate x_j from (5.9). If this generates any inequalities it is equivalent to premultiplying (5.9) by a matrix M^j giving

$$\begin{aligned} (5.10) \quad & M^j(\bar{M}^{j-1} \cdot \ldots \cdot \bar{M}^1)Ax + M^j(\bar{M}^{j-1} \cdot \ldots \cdot \bar{M}^1)By \geqq \\ & M^j(\bar{M}^{j-1} \cdot \ldots \cdot \bar{M}^2 \cdot \bar{M}^1)d \end{aligned}$$

By Theorem 5.1 the rows of the matrix

$$M^j(\bar{M}^{j-1} \cdot \bar{M}^{j-2} \cdot \ldots \cdot \bar{M}^2 \cdot \bar{M}^1)$$

contain a sufficient set of extreme vectors of the convex cone

$$(w_1, w_2, \ldots, w_m) \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} \\ a_{21} & a_{22} & \cdots & a_{2j} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} \end{bmatrix} = 0$$

$$w_i \geqq 0 \;, \qquad\qquad i = 1,2, \ldots, m$$

and therefore we can apply the methods of Theorem 5.2 or Theorem 5.3 or both in order to detect them. If a row of $M^j(\bar{M}^{j-1} \cdot \ldots\ldots \cdot \bar{M}^2 \cdot \bar{M}^1)$ is detected to be *not* a member of a sufficient set of extreme vectors then the corresponding row of M^j is deleted. The result, after all deletions have been made to M^j, is defined to be the matrix $\bar{M}^j$. This completes the elimination of x_j and we are ready to repeat the process on the variable x_{j+1}.

At the end, when x_n has been eliminated (assuming that the method has generated inequalities at every elimination), the rows of the matrix

$$\bar{M}^n \cdot \bar{M}^{n-1} \cdot \ldots\ldots \cdot \bar{M}^2 \cdot \bar{M}^1$$

constitute a sufficient set of extreme vectors of the set C.

Some remarks, mostly concerning computation, are relevant at this point.

(1) Computational experience seems to indicate that it is preferable to apply the method of Theorem 5.3 after the elimination of each variable in order to delete rows of M^j and to use the method of Theorem 5.2 periodically, say after every 10 eliminations or when the number of rows remaining after applying the former method is still large. This is recommended because the criterion of Theorem 5.3 is so easy to apply and seems to eliminate most of the unwanted rows. However the work involved in the method of Theorem 5.2 is also quite small--especially when programmed for a

computer--because binary bits may be used to indicate whether a component of a vector is positive or not and most scientific computers have logical commands, such as AND, OR, EXCLUSIVE OR, which would facilitate the comparisons of the $I(w)$'s .

(2) It is not necessary to calculate the actual values of the elements of the matrix

$$M^j(\bar{M}^{j-1} \cdot \bar{M}^{j-2} \cdot \ldots\ldots \cdot \bar{M}^2 \cdot \bar{M}^1)$$

before rows are deleted from M^j . The criteria of Theorems 5.2 and 5.3 both rely only on whether elements are zero or not zero. The values of the elements should only be calculated for $\bar{M}^j$. In fact, as the rows of M^j are created, a test should be made to see whether they can be deleted immediately. If so then it is unnecessary to keep them of course.

Example

Eliminate x from the system

$$(5.11) \qquad Ax + By \geqq d$$

where

$$A = \begin{bmatrix} 0 & 1 & 1 & 7 \\ 2 & 2 & 0 & -3 \\ 1 & -1 & 0 & -1 \\ -1 & 2 & 0 & 3 \\ -3 & 0 & 0 & \frac{9}{4} \\ -1 & -2 & 0 & 1 \end{bmatrix}$$

(it is not necessary for this purpose to know what B and d are).

(i) Elimination of x_1

Column 1 of A is

$$\begin{bmatrix} 0 \\ 2 \\ 1 \\ -1 \\ -3 \\ -1 \end{bmatrix}$$

and so

$$\bar{M}^{1} = M^{1} = \begin{bmatrix} 1 & & & & & \\ & \frac{1}{2} & & 1 & & \\ & \frac{1}{2} & & & \frac{1}{3} & \\ & \frac{1}{2} & & & & 1 \\ & & 1 & 1 & & \\ & & 1 & & \frac{1}{3} & \\ & & 1 & & & 1 \end{bmatrix}$$

(ii) Elimination of x_2

Premultiplying the second column of A by $\bar{M}^{1}$ we get

$$\bar{M}^{1} \cdot \begin{bmatrix} 1 \\ 2 \\ -1 \\ 2 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ -1 \\ 1 \\ -1 \\ -3 \end{bmatrix}$$

and so

$$M^2 = \begin{bmatrix} 1 & & & 1 & & & \\ 1 & & & & & 1 & \\ 1 & & & & & & \frac{1}{3} \\ & \frac{1}{3} & & 1 & & & \\ & \frac{1}{3} & & & & 1 & \\ & \frac{1}{3} & & & & & \frac{1}{3} \\ & & 1 & 1 & & & \\ & & 1 & & & 1 & \\ & & 1 & & & & \frac{1}{3} \\ & & & 1 & 1 & & \\ & & & & 1 & 1 & \\ & & & & 1 & & \frac{1}{3} \end{bmatrix}$$

The strictly positive elements of $M^2 \cdot \bar{M}^{-1}$ are located as follows:

$$\begin{array}{cl} \begin{bmatrix} + & + & & & & + \\ + & & + & & + & \\ + & & + & & & + \\ & + & & + & & + \\ & + & + & + & + & \\ & + & + & + & & + \\ & + & & & + & + \\ & + & + & & + & \\ & + & + & & + & + \\ & + & + & + & & + \\ & & + & + & + & \\ & & + & + & & + \end{bmatrix} & \begin{array}{l} \\ \\ \\ \\ \leftarrow \text{DELETE} \\ \leftarrow \text{DELETE} \\ \\ \\ \leftarrow \text{DELETE} \\ \leftarrow \text{DELETE} \\ \\ \\ \end{array} \end{array}$$

By Theorem 5.3 we can delete rows 5, 6, 9 and 10 of M^2 because they contain four positive elements. The remaining rows form a sufficient set of extreme vectors as can be seen by applying Theorem 5.2

$$\therefore\ \bar{M}^2 = \begin{bmatrix} 1 & & & 1 & & & \\ 1 & & & & & 1 & \\ 1 & & & & & & \frac{1}{3} \\ & \frac{1}{3} & & 1 & & & \\ & & 1 & 1 & & & \\ & & 1 & & & 1 & \\ & & & & 1 & 1 & \\ & & & & 1 & & \frac{1}{3} \end{bmatrix}$$

and

$$\bar{M}^2 \cdot \bar{M}^1 = \begin{bmatrix} 1 & \frac{1}{2} & & & & 1 \\ 1 & & 1 & & \frac{1}{3} & \\ 1 & & \frac{1}{3} & & & \frac{1}{3} \\ & \frac{2}{3} & & \frac{1}{3} & & 1 \\ & 1 & & & \frac{1}{3} & 1 \\ & \frac{1}{2} & 1 & & \frac{2}{3} & \\ & & 2 & 1 & \frac{1}{3} & \\ & & \frac{4}{3} & 1 & & \frac{1}{3} \end{bmatrix}$$

(iii) Elimination of x_3

Premultiplying the third column of A by $\bar{M}^2 \cdot \bar{M}^1$ we get

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

This is all nonnegative and so we can delete the first three inequalities of the system

$$\bar{M}^2\bar{M}^1Ax + \bar{M}^2\bar{M}^1By \geqq \bar{M}^2\bar{M}^1d$$

$$\therefore\ \bar{M}^3 = M^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & & & & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix}$$

and

$$\bar{M}^3 \cdot \bar{M}^2 \cdot \bar{M}^1 = \begin{bmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 & 1 \\ & 1 & & & \frac{1}{3} & 1 \\ & \frac{1}{2} & 1 & & \frac{2}{3} & \\ & & 2 & 1 & \frac{1}{3} & \\ & & \frac{4}{3} & 1 & & \frac{1}{3} \end{bmatrix}$$

(iv) Elimination of x_4

Premultiplying the fourth column by $\bar{M}^3 \cdot \bar{M}^2 \cdot \bar{M}^1$ we get

$$\begin{bmatrix} 0 \\ -\frac{5}{4} \\ -1 \\ \frac{7}{4} \\ 2 \end{bmatrix}$$

$$\therefore\ M^4 = \begin{bmatrix} 1 & & & & \\ & \frac{4}{5} & & \frac{4}{7} & \\ & & 1 & \frac{4}{7} & \\ & \frac{4}{5} & & & \frac{1}{2} \\ & & 1 & & \frac{1}{2} \end{bmatrix}$$

the positive elements of $M^4(\bar{M}^3\bar{M}^2\bar{M}^1)$ are

$$\begin{array}{cl} \begin{bmatrix} 0 & + & 0 & + & 0 & + \\ 0 & + & + & + & + & + \\ 0 & + & + & + & + & 0 \\ 0 & + & + & + & + & + \\ 0 & + & + & + & + & + \end{bmatrix} & \begin{array}{l} \\ \leftarrow \text{DELETE} \\ \\ \leftarrow \text{DELETE} \\ \leftarrow \text{DELETE} \end{array} \end{array}$$

By Theorem 5.2 we can delete rows 2, 4 and 5

$$\therefore\ \bar{M}^4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{4}{7} & 0 \end{bmatrix}$$

$$\therefore\ \bar{M}^4\bar{M}^3\bar{M}^2\bar{M}^1 = \begin{bmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{3} & 0 & 1 \\ 0 & \frac{1}{2} & \frac{15}{7} & \frac{4}{7} & \frac{6}{7} & 0 \end{bmatrix}$$

and the set Y may be defined by the two inequalities

$$(\bar{M}^4\bar{M}^3\bar{M}^2\bar{M}^1)By \geqq (\bar{M}^4\bar{M}^3\bar{M}^2\bar{M}^1)d$$

The unmodified Fourier-Motzkin would have generated 16 inequalities had it been applied to this problem.

6. COMPUTATIONAL RESULTS

Some tests have been made to compare the method with the Fourier-Motzkin Method and the test problems are given below. Following a short description of each problem there are two rows of figures, e.g.,

12	25	154
12	10	8

The first of these rows gives the number of inequalities generated by the Fourier-Motzkin Method after each elimination, i.e., in this case 12 inequalities are generated when x_1 is eliminated, 25 when x_2 is eliminated, etc. The second row gives the number of inequalities generated by the method described in Section 5. An asterisk, e.g. 1632*, indicates that 1632 is a lower bound and the actual number is almost certain to exceed 1632.

Problem 1: 7 Rows

12	25	154
12	10	8

Problem 2: 15 Rows

14	19	46	425
14	19	26	31

Problem 3: 9 Rows

8	9	14	41	112
8	9	10	9	9

Problem 4: 10 Rows

10	9	17	3	2
10	9	15	3	2

Problem 5: 15 Rows

14	19	39	38	56
14	19	24	23	32

Problem 6: 20 Rows

20	23	21	20	26	53	494
20	22	20	19	23	44	176

and on the next elimination the number of inequalities exceeded 186--the capacity of the machine for this problem.

Problem 7: 20 Rows

19	38	136	177
19	38	90	106

and on the next elimination the number of inequalities exceeded 502--the capacity of the machine for this problem.

Problem 8: 20 Rows

29	28	48	170	1632*	5810*	12391*	27930*
29	28	42	97	195	224	619	288

7. COMPARISON WITH OTHER METHODS

There are other methods for finding extreme vectors or points of a convex polyhedral cone or set, the main ones being Balinski's method [1] and Motzkin's Double Description Method [12] (see Balinski [1] for a discussion of the literature). The remainder of this section will be concerned with a comparative analysis of these two methods and will assume a detailed knowledge of them.

Motzkin's method has been virtually ignored--perhaps because he simply stated a method with some motivation but no proof. The first part of the paper is devoted to a specialization of the Double Description Method (to the problem of finding all solutions of a two-person zero-sum game) which was arrived at independently by Raiffa, Thompson and Thrall who co-authored the paper. The general Double Description Method is given in the last few pages which were written by Motzkin.

Although reference is made to Notes of a Seminar on Linear Programming at the Institute for Numerical Analysis, these contain little additional information. Motzkin assumes that the reader shares his masterly understanding of polyhedral cones--consequently some passages may require careful examination before their meaning becomes clear. Once understood, however, the paper provides profound insight not only into the nature of polyhedral cones and the Double Description Method but also into the method which I have developed independently in this thesis. Motzkin describes three versions of his method:

(i) The general problem $Ax \geqq 0$. This method is very similar to mine in that linear constraints are added one at a time. The main difference is in the formation of a new extreme vector of the cone $A_i x \geqq 0$, $i = 1,2, \ldots, k$ as a convex combination of

two extreme vectors of the cone $A_i x \geq 0$, $i = 1,2, \ldots, k-1$. Motzkin checks first to see whether the latter two vectors are adjacent before generating the new vector. Adjacency is a necessary and sufficient condition that the new vector be extreme. In my method new vectors are formed from all pairs of the old extreme vectors which are on opposite sides of the hyperplane $A_k x = 0$ because we have a "quick and dirty" means, in the form of the criterion of Theorem 5.3, of eliminating nonextreme vectors. Then the criterion of Theorem 5.2 is used to remove any remaining redundant vectors. It is difficult say which, of Motzkin's and my methods, is the more efficient. Possibly an improvement over both would be to order the extreme vectors in some lexicographical manner, dependent upon the position of nonzero elements, so that we could have a "quick and dirty" means of determining adjacency.

(ii) The special case $Ax \geq 0$, $x \geq 0$ where A is "nondegenerate." Motzkin states at the top of Page 70 of his paper [12] that by "nondegenerate" in this context he means that "no $n+1$ of the (inhomogeneous) linear functions vanish at the same point, which is always the case after a small change of the coefficients." By this I am sure he means that no n (or more) of the $m+n$ functions

$$x_j , \qquad j = 1,2, \ldots, n$$

$$A_i x , \qquad i = 1,2, \ldots, m$$

may vanish at the same point, i.e., each solution (x,s) of the system

$$x \geqq 0 \ , \ s \geqq 0 \ ,$$
$$Ax + Is = 0$$

must have more than $((m+n)-n) = m$ nonzeros. Now this condition of nondegeneracy would mean that if we were to apply our method to this problem it would be unnecessary to use the criterion of Theorem 5.2 because the criterion of Theorem 5.3 would eliminate *all* of the redundant vectors; and in fact this version of the Double Description Method is a variant of our method without Theorem 5.2's criterion.

(iii) $Ax \geqq b$, $x \geqq 0$, A nondegenerate and $A > 0$. Although Motzkin doesn't say so the condition $A > 0$ is necessary for the computational method which he gives at the very end of his paper. Without it he could not assume as he does, that

(a) In Step s A , $P\ k\ 0 = 0$ implies that $P_{k\ n+s} > 0$.

(b) $Ax = 0$ has no nonzero solution.

Incidentally, it is interesting to note that in his 1936 thesis Motzkin was within a hair of discovering the relationship between Fourier's method and his own Double Description Method but made an error in the logic of Section 86 (assuming that $(vG_1)p_1$ is identical to $v(G_1p_1)$) . We can see now that the Double Description Method and my method both become, in essence, Fourier's method if new vectors generated by two old *non*adjacent vectors are not eliminated at each iteration.

Balinski's method is not strictly comparable with our method since it is designed to find only the vertices (i.e., *non*homogeneous extreme vectors) of the system $Ax \geqq b$ where the columns of A are linearly independent. By putting this system in the form

$$
(7.1) \quad \begin{aligned} & Ax' - Ax'' - Is - bz = 0 \\ & x' , x'' , s , z \geq 0 \end{aligned}
$$

we see that our method could solve Balinski's problem (merely select those extreme vectors in which $z > 0$) . However, Balinski's method does not find homogeneous solutions and so it does not directly solve our problem. I feel it can be modified to do this but has not been to date. Balinski has written a computing code for the special problem

$$
Ax \geq b , x \geq 0
$$

and an attempt was made to compare efficiencies with actual problems using a computer. The tests were inconclusive as we will show below.

Test Problem 1

When stated in the form (7.1) the A matrix of this problem had 15 columns and 4 rows. The CDC 6400 computer was used and it required 2.359 seconds to produce the 248 extreme vectors using my code. 82 of these were vertices. Balinski's code on the same machine required 5.655 seconds and 1503 points (iterations) were examined. However, there was some degeneracy which caused some vertices to be generated many times by Balinski's code. Altogether his code generated 126 solutions which it classed as "vertices;" 44 out of the 126 were duplicates. So although Balinski's code required more time it was in effect doing unnecessary work. The problem of degeneracy is a significant one for Balinski's method because most practical problems are degenerate and, although this can be eliminated by perturbing the coefficients slightly, my experience has been that this increases the number of extreme vectors (and hence vertices) enormously.

Test Problem 2

A had 11 columns and 8 rows. My code required 20 seconds to generate the 336 extreme vectors of which 90 were vertices. Balinski's code required 285 seconds and 45,370 iterations to solve the problem. However, again there was much duplication of vertices--but not caused by degeneracy this time. Since theoretically vertices cannot be repeated I conclude that there is an error in Balinski's code.

In conclusion, the indication--*not* definite--is that Balinski's method is less efficient for generating *all* vertices. However, there are certain problems in which one is looking for *one* vertex with certain characteristics and for this Balinski's method seems well suited because it passes progressively from one vertex to another.

8. APPLICATIONS

(i) Linear Programming With Varying Cost Coefficients

Consider the problem

$$(8.1)\qquad \begin{array}{ll} \text{Minimize} & cx \\ \text{Subject to} & Ax = b \;,\; x \geqq 0 \end{array}$$

which is to be solved many times (say every day) with fixed A and b but varying c. We can find the set of all extreme points of $Ax = b$, $x \geqq 0$ by finding all extreme vectors of the cone

$$x \geqq 0 \;,\; \lambda \geqq 0 \;,$$
$$Ax - b\lambda = 0 \;,$$

using our method, and discarding those extreme vectors for which $\lambda = 0$. Then the linear program may be solved by simply choosing the extreme vector x which yields the smallest cx. This method is suited to a problem of this sort because, although considerable work is involved in generating all extreme points, this need only be done once.

(ii) Linear Programming With Varying Right-Hand Sides

Consider the problem

$$\begin{array}{ll} \text{Minimize} & cx \\ \text{Subject to} & Ax \geqq b \;,\; x \geqq 0 \end{array}$$

but this time we will assume that A and c remain fixed while b varies. Then by eliminating x from the system

$$\begin{array}{rl} Ix & \geqq 0 \;, \\ Ax & \geqq b \;, \\ -cx + z & \geqq 0 \;, \end{array}$$

we can solve all subsequent problems very simply. To be specific we need to do the following:

(a) Obtain all extreme vectors (s^i, w^i, λ^i) , $i = 1,2, \ldots$ of the cone

$$sI + wA - \lambda c = 0 ,$$
$$s , w , \lambda \geqq 0$$

for which $\lambda^i > 0$.

(b) For each i compute and store $\frac{1}{\lambda^i} \cdot w^i$, $i = 1,2, \ldots$

Then for each b we simply form the inequalities

$$z \geqq \frac{1}{\lambda^i} \cdot w^i b , \qquad i = 1,2, \ldots$$

and choose the i which gives the largest value of $\frac{1}{\lambda^i} \cdot w^i b$. The corresponding value of x may be formed using the method of (iii)c below.

<u>(iii) Finding All Solutions Of A Linear Program</u>

This can be achieved in at least three ways

(a) By use of (i) above and by selection of all extreme vectors x which yield the minimum value of cx .

(b) Solve the problem first by the simplex method giving, say, $\text{Min } cx = z_o$ and then find all extreme points of the polyhedral set

$$x \geqq 0 ,$$
$$Ax = b ,$$
$$cx = z_o$$

(c) Eliminate x from the system

$$
(8.2) \qquad \begin{aligned} Ix &\geqq 0 \\ Ax &\geqq b \\ -cx + z &\geqq 0 \end{aligned}
$$

yielding a system of inequalities in the variable z, of the form

$$
(8.3) \qquad z \geqq wb
$$

where w is an extreme point of the polyhedral set

$$
(8.4) \qquad \begin{aligned} -c + wA &\leqq 0 \\ w &\geqq 0 \end{aligned}
$$

(i.e., if (w^o, w, s) is an extreme vector of the polyhedral cone

$$
\begin{gathered} -w^o c + wA + sI = 0 , \\ w^o \geqq 0 , \ w \geqq 0 , \ s \geqq 0 \end{gathered}
$$

we select only those extreme vectors for which $w^o > 0$ and insist, for them, that w^o be equal to 1). Now let z_o be the minimum value of z. It will satisfy some of the inequalities (8.3) as equalities. Taking each of these equalities

$$
z_o = w'b
$$

in turn we notice that it is a *positive* linear combination of some of the equalities of (8.2) which must therefore also be satisfied as equalities by the optimal value of x and z_o. Now (w^o, w, s) is an extreme point and therefore the subsystem of equations contains no redundant row, is nonsingular and may be solved by inverting its matrix of coefficients.

Incidentally, logic similar to the above may be used to derive the Duality Theorem of Linear Programming.

(iv) Finding All Solutions Of Bimatrix Games

O. L. Mangasarian [10] has given a method for finding the set of all equilibrium points of bimatrix game. His method uses Balinski's algorithm for finding all vertices of a convex polyhedral set. Our method may be substituted for Balinski's in this context.

(v) Bilinear Programming

As Mangasarian and Stone showed [9], finding an equilibrium point of a bimatrix game is equivalent to finding a solution of a certain type of bilinear program. The general bilinear programming problem may be stated as

$$\begin{aligned} &\text{Minimize} \quad xCy + px + qy \\ &\text{Subject to} \quad Ax = b \\ &\qquad By = d \\ &\qquad x \geqq 0 \;,\; y \geqq 0 \end{aligned} \tag{8.5}$$

and it may be solved by two methods at least:

(a) As is well known, there is a solution (x^o, y^o) of (8.5) such that x^o is an extreme point of

$$Ax = b \;,\; x \geqq 0 \tag{8.6}$$

and y^o is an extreme point of

$$By = d \;,\; y \geqq 0 \;. \tag{8.7}$$

And so we can take all possible pairs of extreme points, one of (8.6) and the other of (8.7), and choose the pair(s) which give the

minimum value of $xCy + px + qy$. This is basically what Mangasarian did (see (iii) above).

(b) (K Murty is responsible for the following idea:) In (8.5) let us assume for the moment that we have fixed x and we want to choose the optimal y . (8.5) then becomes a linear program whose dual is:

$$\begin{aligned} &\underset{u}{\text{Maximize}} \quad z \\ &\text{Subject to} \quad ud - z = 0 \\ &\qquad\qquad uB - xC \leqq 0 \\ &\qquad\qquad\qquad Ax = b \\ &\qquad\qquad\qquad\; x \geqq 0 \,. \end{aligned} \tag{8.8}$$

Now for fixed x , the system (8.8) defines a *convex* set in u and z which means that $\underset{u}{\text{Max}}\, z$ is a *concave* function of x . The problem (8.5) therefore becomes $\underset{x}{\text{Minimize}}\ \{\underset{u}{\text{Max}}\, ud \mid uB - xC \leqq 0$, $Ax = b$, $x \geqq 0\}$ that is to say, a problem of *minimizing* a *concave* function, $(\underset{u}{\text{Max}}\, ud \mid uB - xC \leqq 0)$, over a convex set $(Ax = b$, $x \geqq 0)$. One approach to this problem is to first eliminate u from the system (8.8) giving a system in z and x ; then find the extreme points of $Ax = b$, $x \geqq 0$ substituting each time to find the minimum possible value of z .

(vi) Minimizing A Concave Function Over A Convex Polyhedral Set

Mangasarian in his forthcoming book on Nonlinear Programming gives the following theorem (Section 5.1, Theorem 8):

Let R be convex and let $g(x)$ be concave on R . If $g(x)$ is not constant on R then no interior point of R can solve the problem

$$\underset{x \in R}{\text{Minimize}} \quad g(x) \,.$$

This implies that if R is defined by *linear* constraints and if $g(x)$ is not constant then the minimal value is achieved at an extreme point. Therefore, our method can again be used to generate all extreme points and thereby solve this problem.

REFERENCES

[1] Balinski, M., "An Algorithm for Finding all Vertices of Convex Polyhedral Sets," J.Soc.Indust.Appl.Math., Vol. 9, pp. 72-88, (1961).

[2] Chernikov, S. N., "Contraction of Systems of Linear Inequalities," Dokl.Akad.Nauk., (SSSR), Vol. 131, pp. 518-521, (1960). Translation in Soviet Mathematics, Vol. 1, No. 2, pp. 296-299.

[3] Chernikov, S. N., "The Solution of Linear Programming Problems by Elimination of Unknowns," Dokl.Akad.Nauk., (SSSR), Vol. 139, pp. 1314-1317, (1961). Translation in Soviet Mathematics, Vol. 2, No. 4, pp. 1099-1103.

[4] Chernikov, S. N., "Contraction of Finite Systems of Linear Inequalities," Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, (USSR), Vol. 5, No. 1, pp. 3-20, (1965). Abstracted in Mathematical Reviews, Vol. 32 and in Cybernetics Abstracts, No. 7, T-153, (1965).

[5] Dantzig, G. B., LINEAR PROGRAMMING AND EXTENSIONS, Princeton University Press, Princeton, New Jersey, (1963).

[6] Farkas, J., "Über die Theorie der Einfachen Ungleichungen," J. Reine Angew Math., Vol. 124, pp. 1-24, (1902). (Or see References 5 or 8.)

[7] Fourier, Jean Baptiste Joseph, "Solution d'une Question Particulière du Calcul des Inégalités," (1826), and extracts from "Histoire de l'Académie," (1823, 1824), Oeuvres II, pp. 317-328.

[8] Gale, David, "The Basic Theorems of Real Linear Equations, Inequalities, Linear Programming and Game Theory," Naval Res.Logist.Quart., Vol. 3, No. 3, pp. 193-200, (September 1956).

[9] Mangasarian, O. L. and H. Stone, "Two-Person Nonzero-Sum Games and Quadratic Programming," J.Math.Anal. and Appl., Vol. 9, No. 3, (December 1964).

[10] Mangasarian, O. L., "Equilibrium Points of Bimatrix Games," J.Soc. Indust.Appl.Math., Vol. 12, No. 4, (December 1964).

[11] Motzkin, T. S., "Beiträge zur Theorie der Linearen Ungleichungen," Jerusalem, (1936), (Doctoral Thesis, University of Zurich). (Unapproved translation by D. R. Fulkerson, "The Theory of Linear Inequalities," RAND Corporation, Report T-22, (March 1952).)

[12] Motzkin, T. S., H. Raiffa, G. L. Thompson and R. M. Thrall, "The Double Description Method," in H. W. Kuhn and A. W. Tucker (eds.), CONTRIBUTIONS TO THE THEORY OF GAMES, Vol. II, Annals of Mathematics Study, No. 28, Princeton University Press, Princeton, New Jersey, pp. 51-73, (1953).

[13] Motzkin, T. S., Lecture notes for a seminar on Linear Programming, Institute for Numerical Analysis, University of California, Los Angeles.

[14] Wets, Roger J.B. and Christoph Witzgall, "Algorithms for Frames and Lineality Spaces of Cones," Journal of Research of National Bureau of Standards, B--Mathematics and Mathematical Physics, B, Vol. 71, No. 1, (January-March 1967).

[15] Wets, Roger J.B., "Programming under Uncertainty: The Solution Set," Boeing Document D1-82-0464, Boeing Scientific Research Laboratories, Seattle, Washington, (August 1965).

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) University of California, Berkeley	2a REPORT SECURITY CLASSIFICATION Unclassified 2b GROUP

3 REPORT TITLE

PROJECTIONS OF CONVEX POLYHEDRAL SETS

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)

Research Report

5 AUTHOR(S) (Last name, first name, initial)

KOHLER, David A.

6 REPORT DATE	7a TOTAL NO OF PAGES	7b NO OF REFS
August 1967	48	15

8a CONTRACT OR GRANT NO Nonr-222(83) b PROJECT NO NR 047 033 c Research Project No.: RR 003 07 01 d	9a ORIGINATOR'S REPORT NUMBER(S) ORC 67-29 9b OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10 AVAILABILITY/LIMITATION NOTICES

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED.

11 SUPPLEMENTARY NOTES Also supported by the National Science Foundation under Grant GK-1684.	12 SPONSORING MILITARY ACTIVITY MATHEMATICAL SCIENCE DIVISION

13 ABSTRACT

The main problem considered is: Given a set of linear inequalities

$$Ax + By \geq d \,, \tag{1.1}$$

which defines a set of $(x;y)$, find and concisely define a set Y of y such that if $(x;y)$ solves (1.1) then y belongs to Y and, conversely, if y belongs to Y then there exists an x such that $(x;y)$ solves (1.1).

The solution to this problem involves finding the set of all extreme rays of the convex cone

$$wA = 0 \,, \; w \geq 0$$

and a method is given for this. The method is compared with other methods for finding extreme rays and points and finally some practical applications are given.

DD FORM 1 JAN 64 1473

Unclassified

Security Classification

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Projections Linear Inequalities Extreme Points						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8a. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through

______________________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through

______________________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through

______________________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.